Modelos no paramétricos y de regresión/Estadística II Semestre 2018-1

Tarea 4

Fecha de entrega: 18 de septiembre

Notación:

- $\mathbf{I}_n$: matriz identidad de dimensión n.
- $\mathbf{1}_n$: vector de unos de dimensión n.
- $\mathbf{J}_n$: matriz de unos de dimensión $n \times n$.

1. Considerar las siguientes matrices

$$\mathbf{P} = \begin{pmatrix} p^2 & 2pq & q^2 \\ p^2 & 2pq & q^2 \\ p^2 & 2pq & q^2 \end{pmatrix} \qquad \text{y} \qquad \mathbf{T} = \begin{pmatrix} p & q & 0 \\ \frac{1}{2}p & \frac{1}{2} & \frac{1}{2}q \\ 0 & p & q \end{pmatrix}$$

con $p + q = 1$ y $p, q > 0$. Mostrar que

a) $\mathbf{P} = \mathbf{1}_3 \begin{pmatrix} p^2 & 2pq & q^2 \end{pmatrix}$,

b) $\mathbf{P}\mathbf{1}_3 = \mathbf{1}_3$,

c) $\mathbf{T}\mathbf{1}_3 = \mathbf{1}_3$,

d) $\mathbf{TP} = \mathbf{PT} = \mathbf{P}$,

e) $\mathbf{T}^2 = \frac{1}{2}(\mathbf{P} + \mathbf{T})$,

f) $\mathbf{P}^2 = \mathbf{P}$.

2. *a)* Mostrar que $\left(\mathbf{I}_n - \frac{1}{n}\mathbf{J}_n\right)$ es idempotente.

b) Si p, q r y s son escalares, mostrar que

$$(p\mathbf{I}_n + q\mathbf{J}_n)(r\mathbf{I}_n + s\mathbf{J}_n) = pr\mathbf{I}_n + (ps + qr + nqs)\mathbf{J}_n.$$

c) Si p y q son escalares, con $p \neq 0$ y $p + qn \neq 0$, simplificar

$$(p\mathbf{I}_n + q\mathbf{J}_n)\frac{1}{p}\left(\mathbf{I}_n - \frac{q}{p + qn}\mathbf{J}_n\right).$$

3. Suponer que se tienen n observaciones de p variables representadas en la siguiente matriz

$$\mathbf{X} = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{pmatrix}$$

a) Mostrar que $\bar{\mathbf{x}}_n = \frac{1}{n}\mathbf{X}'\mathbf{1}$. Donde $\bar{\mathbf{x}}_n$ es un vector de dimensión n y con i-ésima entrada

$$\bar{x}_i = \frac{1}{n}\sum_{k=1}^{n} x_{ki}$$

b) Mostrar que $\mathbf{S}_n = \frac{1}{n-1}\mathbf{X}'\left(\mathbf{I}_n - \frac{1}{n}\mathbf{J}_n\right)\mathbf{X}$. Donde $\mathbf{S}_n$ es una matriz cuadrada de dimensión n con ij-ésima entrada

$$\frac{1}{n-1}\sum_{k=1}^{n}(x_{ki} - \bar{x}_i)(x_{kj} - \bar{x}_j)$$

4. Considerar las siguientes matrices

$$\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \qquad \text{y} \qquad \mathbf{X} = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}$$

Calcular las siguientes expresiones

a) $\mathbf{X}'\mathbf{X}$.

b) $|\mathbf{X}'\mathbf{X}|$.

c) $(\mathbf{X}'\mathbf{X})^{-1}$ (¿qué se debe cumplir para que tal inversa exista?)

d) $(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$

5. Si $\mathbf{X} = (X_1, X_2, X_3)' \sim N_3(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ con

$$\boldsymbol{\mu} = \begin{pmatrix} -3 \\ 1 \\ 4 \end{pmatrix} \qquad \text{y} \qquad \boldsymbol{\Sigma} = \begin{pmatrix} 1 & -2 & 0 \\ -2 & 5 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

Resonder a lo siguiente.

a) ¿Cuáles de los siguientes pares de variables aleatorias son independientes? Explicar por qué: *i)* X_1 y X_2,*ii)* X_2 y X_3,*iii)* X_1 y $X_1 + 3X_2 - 2X_3$.

b) Indicar cuál es la distribución de las siguientes variables o vectores aleatorios: *i)* X_1, *ii)* X_3, *iii)* (X_1, X_2).

c) Sea la matriz

$$\mathbf{A} = \begin{pmatrix} \frac{1}{2} & -\frac{1}{4} & -\frac{1}{4} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$$

¿Cuál es la distribución del vector $\mathbf{AX}$?

⋆ Opcional: Mostrar que si $a \neq 0$ y $a + bn \neq 0$, entonces

$$(a\mathbf{I}_n + b\mathbf{J}_n)^{-1} = \frac{1}{a}\left(\mathbf{I}_n - \frac{b}{a+bn}\mathbf{J}_n\right).$$